

M. Walicka

Re-run



RAW SEQUENCE LISTING

DATE: 01/09/2003

PATENT APPLICATION: US/09/308,207

TIME: 11:14:33

Input Set : N:\vernette\US09308207.raw

Output Set: N:\CRF4\01092003\I308207.raw

SEQUENCE LISTING

1 (1) GENERAL INFORMATION:

2 (i) APPLICANT: MARIA DIAZ-TORRES ET AL.

3 (ii) TITLE OF INVENTION: METHOD FOR THE RECOMBINANT

4 PRODUCTION OF 1,3 PROPANEDIOL

5 (iii) NUMBER OF SEQUENCES: 68

6 (iv) CORRESPONDENCE ADDRESS:

7 (A) ADDRESSEE: Genencor International, Inc.

8 (B) STREET: 4 Cambridge Place

9 1870 South Winton road

10 (C) CITY: Rochester

11 (D) STATE: NY

12 (E) COUNTRY: U.S.A

13 (F) ZIP: 14618

14 (v) COMPUTER READABLE FORM:

15 (A) MEDIUM TYPE: Diskette

16 (B) COMPUTER: IBM Compatible

17 (C) OPERATING SYSTEM: Windows

18 (D) SOFTWARE: FastSEQ for Windows Version 2.0b

19 (vi) CURRENT APPLICATION DATA:

C--> 20 (A) APPLICATION NUMBER: US/09/308,207

C--> 21 (B) FILING DATE: 13-May-1999

22 (C) CLASSIFICATION:

23 (vii) PRIOR APPLICATION DATA:

24 (A) APPLICATION NUMBER: 60/030,601

25 (B) FILING DATE: 13-NOV-1996

26 (viii) ATTORNEY/AGENT INFORMATION:

27 (A) NAME: Glaister, Debra

28 (B) REGISTRATION NUMBER: 33,888

29 (C) REFERENCE/DOCKET NUMBER: GC 369-2

30 (ix) TELECOMMUNICATION INFORMATION:

31 (A) TELEPHONE: 650-864-7620

32 (B) TELEFAX: 650-845-6504

33 (C) TELEX:

34 (2) INFORMATION FOR SEQ ID NO: 1:

35 (i) SEQUENCE CHARACTERISTICS:

36 (A) LENGTH: 1668 base pairs

37 (B) TYPE: nucleic acid

38 (C) STRANDEDNESS: single

39 (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: DNA (genomic)

41 (iii) HYPOTHETICAL: NO

42 (iv) ANTI-SENSE: NO

ENTERED

RAW SEQUENCE LISTING

DATE: 01/09/2003

PATENT APPLICATION: US/09/308,207

TIME: 11:14:33

Input Set : N:\vernette\US09308207.raw

Output Set: N:\CRF4\01092003\I308207.raw

```

43 (vi) ORIGINAL SOURCE:
44 (A) ORGANISM: DHAB1
45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
46 ATGAAAAGAT CAAAACGATT TGCAGTACTG GCCCAGCGCC CCGTCAATCA GGACGGGCTG 60
47 ATTGGCGAGT GGCCTGAAGA GGGGCTGATC GCCATGGACA GCCCCTTTGA CCCGGTCTCT 120
48 TCAGTAAAAG TGGACAACGG TCTGATCGTC GAACTGGACG GCAAACGCCG GGACCAGTTT 180
49 GACATGATCG ACCGATTTAT CGCCGATTAC GCGATCAACG TTGAGCGCAC AGAGCAGGCA 240
50 ATGCGCCTGG AGGCGGTGGA AATAGCCCGT ATGCTGGTGG ATATTACAGT CAGCCGGGAG 300
51 GAGATCATTG CCATCACTAC CGCCATCACG CCGGCCAAAAG CGGTCGAGGT GATGGCGCAG 360
52 ATGAACGTGG TGGAGATGAT GATGGCGCTG CAGAAGATGC GTGCCGCCG GACCCCTCC 420
53 AACCAGTGCC ACGTCACCA TCTCAAAGAT AATCCGGTGC AGATTGCCGC TGACGCCGCC 480
54 GAGGCCGGA TCCGCGGCTT CTCAGAACAG GAGACCACGG TCGGTATCGC GCGCTACGCG 540
55 CCGTTTAACG CCCTGGCGCT GTTGGTCGGT TCGCAGTGGC GCCGCCCCG CGTGTTGACG 600
56 CAGTGCTCGG TGGAAGAGGC CACCGAGCTG GAGCTGGGCA TCGTGCGCTT AACCAGCTAC 660
57 GCCGAGACGG TGTCGGTCTA CGGCACCGAA GCGGTATTTA CCGACGGCGA TGATACGCCG 720
58 TGGTCAAAGG CGTTCCTCGC CTCGGCCTAC GCCTCCCGCG GGTGAAAAAT GCGCTACACC 780
59 TCCGGCACCG GATCCGAAGC GCTGATGGGC TATTCGAGAG GCAAGTCGAT GCTCTACCTC 840
60 GAATCGCGCT GCATCTTCAT TACTAAAGGC GCCGGGGTTC AGGGACTGCA AAACGGCGCG 900
61 GTGAGCTGTA TCGGCATGAC CGGCGTGTG CCGTCGGGCA TTCGGGCGGT GCTGGCGGAA 960
62 AACCTGATCG CCTCTATGCT CGACCTCGAA GTGGCGTCCG CCAACGACCA GACTTTCTCC 1020
63 CACTCGGATA TTCGCCGCAC CGCGCGCACC CTGATGCAGA TGCTGCCGGG CACCGACTTT 1080
64 ATTTTCTCCG GCTACAGCGC GGTGCCGAAC TACGACAACA TGTTCCGCCG CTCGAACTTC 1140
65 GATGCGGAAG ATTTTGTATGA TTACAACATC CTGCAGCGTG ACCTGATGGT TGACGGCGGC 1200
66 CTGCGTCCGG TGACCGAGGC GAAAACCATT GCCATTCGCC AGAAAGCGGC GCGGGCGATC 1260
67 CAGGCGGTTT TCCGCGAGCT GGGGCTGCCG CCAATCGCCG ACGAGGAGGT GGAGGCCGCC 1320
68 ACCTACGCGC ACGGCAGCAA CGAGATGCCG CCGCGTAACG TGGTGGAGGA TCTGAGTGCG 1380
69 GTGGAAGAGA TGATGAAGCG CAACATCACC GGCTCGATA TTGTCGGCGC GCTGAGCCGC 1440
70 AGCGGCTTTG AGGATATCGC CAGCAATATT CTCAATATGC TGCGCCAGCG GGTCACCGCG 1500
71 GATTACCTGC AGACCTCGGC CATTCTCGAT CGGCAGTTTC AGGTGGTGAG TGCGGTCAAC 1560
72 GACATCAATG ACTATCAGGG GCCGGGCACC GGCTATCGCA TCTCTGCCGA ACGCTGGGCG 1620
73 GAGATCAAAA ATATTCCGGG CGTGGTTTCAG CCCGACACCA TTGAATAA 1668
75 (2) INFORMATION FOR SEQ ID NO: 2:
76 (i) SEQUENCE CHARACTERISTICS:
77 (A) LENGTH: 585 base pairs
78 (B) TYPE: nucleic acid
79 (C) STRANDEDNESS: single
80 (D) TOPOLOGY: linear
81 (ii) MOLECULE TYPE: DNA (genomic)
82 (vi) ORIGINAL SOURCE:
83 (A) ORGANISM: DHAB2
84 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
85 GTGCAACAGA CAACCCAAAT TCAGCCCTCT TTTACCCTGA AAACCCGCGA GGGCGGGGTA 60
86 GCTTCTGCCG ATGAACGCGC CGATGAAGTG GTGATCGGCG TCGGCCCTGC CTCGATAAAA 120
87 CACCAAGCATC AACTCTGAT CGATACGCC CATGGCCGCA TCCTCAAAGA GCTGATTGCC 180
88 GGGGTGGAAG AAGAGGGGCT TCACGCCCCG GTGGTGCACA TTCTGCGCAC GTCCGACGTC 240
89 TCCTTTATGG CCTGGGATGC GGCCAACCTG AGCGGCTCGG GGATCGGCAT CGGTATCCAG 300
90 TCGAAGGGGA CCACGGTCAT CCATCAGCGC GATCTGCTGC CGCTCAGCAA CCTGGAGCTG 360
91 TTCTCCAGG CGCCGCTGCT GACGCTGGAG ACCTACCGGC AGATTGGCAA AAACGCTGCG 420
92 CGCTATGCGC GCAAAGAGTC ACCTTCGCCG GTGCCGGTGG TGAACGATCA GATGGTGCGG 480

```

RAW SEQUENCE LISTING

DATE: 01/09/2003

PATENT APPLICATION: US/09/308,207

TIME: 11:14:33

Input Set : N:\vernette\US09308207.raw

Output Set: N:\CRF4\01092003\I308207.raw

```

93      CCGAAATTTA TGGCCAAAGC CGCGCTATTT CATATCAAAG AGACCAAACA TGTGGTGCAG      540
94      GACGCCGAGC CCGTCACCCT GCACATCGAC TTAGTAAGGG AGTGA                      585
96 (2) INFORMATION FOR SEQ ID NO: 3:
97      (i) SEQUENCE CHARACTERISTICS:
98          (A) LENGTH: 426 base pairs
99          (B) TYPE: nucleic acid
100         (C) STRANDEDNESS: single
101         (D) TOPOLOGY: linear
102      (ii) MOLECULE TYPE: DNA (genomic)
103      (vi) ORIGINAL SOURCE:
104          (A) ORGANISM: DHAB3
105      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
106      ATGAGCGAGA AAACCATGCG CGTGCAGGAT TATCCGTTAG CCACCCGCTG CCCGGAGCAT      60
107      ATCCTGACGC CTACCGGCAA ACCATTGACC GATATTACCC TCGAGAAGGT GCTCTCTGGC      120
108      GAGGTGGGCC CGCAGGATGT GCGGATCTCC CGCCAGACCC TTGAGTACCA GGCGCAGATT      180
109      GCCGAGCAGA TGCAGCGCCA TGCGGTGGCG CGCAATTTCC GCCGCGCGGC GGAGCTTATC      240
110      GCCATTCTTG ACGAGCGCAT TCTGGCTATC TATAACGCGC TGCGCCCGTT CCGCTCCTCG      300
111      CAGGCGGAGC TGCTGGCGAT CGCCGACGAG CTGGAGCACA CCTGGCATGC GACAGTGAAT      360
112      GCCGCCTTTG TCCGGGAGTC GCGGGAAGTG TATCAGCAGC GGCATAAGCT GCGTAAAGGA      420
113      AGCTAA                                                                426
115 (2) INFORMATION FOR SEQ ID NO: 4:
116      (i) SEQUENCE CHARACTERISTICS:
117          (A) LENGTH: 1164 base pairs
118          (B) TYPE: nucleic acid
119          (C) STRANDEDNESS: single
120          (D) TOPOLOGY: linear
121      (ii) MOLECULE TYPE: DNA (genomic)
122      (vi) ORIGINAL SOURCE:
123          (A) ORGANISM: DHAT
124      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
125      ATGAGCTATC GTATGTTTGA TTATCTGGTG CCAAACGTTA ACTTTTTTTGG CCCCAACGCC      60
126      ATTTCCGTAG TCGGCGAACG CTGCCAGCTG CTGGGGGGGA AAAAAGCCCT GCTGGTCACC      120
127      GACAAAGGCC TCGGGGCAAT TAAAGATGGC GCGGTGGACA AAACCTTGCA TTATCTGCGG      180
128      GAGGCCGGGA TCGAGGTGGC GATCTTTGAC GGCGTCGAGC CGAACCCGAA AGACACCAAC      240
129      GTGCGCGACG GCCTCGCCGT GTTTCGCCGC GAACAGTGCG ACATCATCGT CACCGTGGGC      300
130      GCGGCGAGCC CGCACGATTG CGGCAAAGGC ATCGGCATCG CCGCCACCCA TGAGGGCGAT      360
131      CTGTACCAGT ATGCCGGAAT CGAGACCCTG ACCAACCCGC TGCCGCCTAT CGTCGCGGTC      420
132      AATACCACCG CCGGCACCGC CAGCGAGGTC ACCCGCCACT GCGTCCTGAC CAACACCGAA      480
133      ACCAAAGTGA AGTTTGTGAT CGTCAGCTGG CGAAACTGCG CGTCGGTCTC TATCAACGAT      540
134      CCACTGCTGA TGATCGGTAA ACCGGCCGCC CTGACCGCGG CGACCGGGAT GGATGCCCTG      600
135      ACCCAGCCCG TAGAGGCCTA TATCTCCAAA GACGCTAACC CCGTGACGGA CGCCGCCGCC      660
136      ATGCAGGCGA TCCGCCTCAT CGCCCACAAC CTGCGCCAGG CCGTGGCCCT CGGCAGCAAT      720
137      CTGCAGGCGC GGGAAAACAT GGCTATGCT TCTCTGCTGG CCGGGATGGC TTTCAATAAC      780
138      GCCAACCTCG GCTACGTGCA CGCCATGGCG CACCAGCTGG GCGGCCTGTA CGACATGCCG      840
139      CACGGCGTGG CCAACGCTGT CTGCTGCCG CATGTGGCGC GCTACAACCT GATCGCCAAC      900
140      CCGGAGAAAT TCGCCGATAT CGCTGAACTG ATGGGCGAAA ATATCACCGG ACTGTCCACT      960
141      CTCGACGCGG CGGAAAAAGC CATCGCCGCT ATCACGCGTC TGTCGATGGA TATCGGTATT      1020
142      CCGCAGCATC TGCGCGATCT GGGGGTAAAA GAGGCCGACT TCCCCTACAT GGCGGAGATG      1080
143      GCTCTAAAAG ACGGCAATGC GTTCTCGAAC CCGCGTAAAG GCAACGAGCA GGAGATTGCC      1140

```

RAW SEQUENCE LISTING

DATE: 01/09/2003

PATENT APPLICATION: US/09/308,207

TIME: 11:14:33

Input Set : N:\vernette\US09308207.raw

Output Set: N:\CRF4\01092003\I308207.raw

```

144      GCGATTTTCC GCCAGGCATT CTGA                                     1164
146 (2) INFORMATION FOR SEQ ID NO: 5:
147      (i) SEQUENCE CHARACTERISTICS:
148          (A) LENGTH: 1380 base pairs
149          (B) TYPE: nucleic acid
150          (C) STRANDEDNESS: single
151          (D) TOPOLOGY: linear
152      (ii) MOLECULE TYPE: DNA (genomic)
153      (vi) ORIGINAL SOURCE:
154          (A) ORGANISM: GPD1
155      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
156      CTTTAATTTT CTTTATCTT ACTCTCTAC ATAAGACATC AAGAAACAAT TGTATATTGT      60
157      ACACCCCCC CCTCCACAAA CACAAATATT GATAATATAA AGATGTCTGC TGCTGCTGAT      120
158      AGATTAAACT TAACTTCCGG CCACTTGAAT GCTGGTAGAA AGAGAAGTTC CTCTTCTGTT      180
159      TCTTTGAAGG CTGCCGAAAA GCCTTTCAAG GTTACTGTGA TTGGATCTGG TAACTGGGGT      240
160      ACTACTATTG CCAAGGTGGT TGCCGAAAAAT TGTAAGGGAT ACCCAGAAGT TTTGCTCCA      300
161      ATAGTACAAA TGTGGGTGTT CGAAGAAGAG ATCAATGGTG AAAAAATTGAC TGAAATCATA      360
162      AATACTAGAC ATCAAAACGT GAAATACTTG CCTGGCATCA CTCTACCCGA CAATTTGGTT      420
163      GCTAATCCAG ACTTGATTGA TTCAGTCAAG GATGTCGACA TCATCGTTTT CAACATTCCA      480
164      CATCAATTTT TGCCCCGTAT CTGTAGCCAA TTGAAAGGTC ATGTTGATTC ACACGTCAGA      540
165      GCTATCTCCT GTCTAAAGGG TTTTGAAGTT GGTGCTAAAG GTGTCCAATT GCTATCCTCT      600
166      TACATCACTG AGGAACTAGG TATTCAATGT GGTGCTCTAT CTGGTGCTAA CATTGCCACC      660
167      GAAGTCGCTC AAGAACACTG GTCTGAAACA ACAGTTGCTT ACCACATTCC AAAGGATTTC      720
168      AGAGGCGAGG GCAAGGACGT CGACCATAAG GTTCTAAAGG CCTTGTTCCA CAGACCTTAC      780
169      TTCCACGTTA GTGTCATCGA AGATGTTGCT GGTATCTCCA TCTGTGGTGC TTTGAAGAAC      840
170      GTTGTGCTT TAGGTTGTGG TTTCGTCGAA GGTCTAGGCT GGGGTAACAA CGCTTCTGCT      900
171      GCCATCCAAA GAGTCGGTTT GGGTGAGATC ATCAGATTCG GTCAAATGTT TTTCCAGAA      960
172      TCTAGAGAAG AAACATACTA CCAAGAGTCT GCTGGTGTTG CTGATTGAT CACCACCTGC      1020
173      GCTGGTGGTA GAAACGTCAA GGTGCTAGG CTAATGGCTA CTTCTGGTAA GGACGCTGG      1080
174      GAATGTGAAA AGGAGTTGTT GAATGGCCAA TCCGCTCAAG GTTTAATTAC CTGCAAAGAA      1140
175      GTTCACGAAT GGTGGAAC ATGTGGCTCT GTCGAAGACT TCCCATTATT TGAAGCCGTA      1200
176      TACCAATCG TTTACAACAA CTACCCAATG AAGAACCTGC CGGACATGAT TGAAGAATTA      1260
177      GATCTACATG AAGATTAGAT TTATTGGAGA AAGATAACAT ATCATACTTC CCCCCTTTT      1320
178      TTCGAGGCTC TTCTATATCA TATTCATAAA TTAGCATTAT GTCATTCTC ATAACTACTT      1380
180 (2) INFORMATION FOR SEQ ID NO: 6:
181      (i) SEQUENCE CHARACTERISTICS:
182          (A) LENGTH: 2946 base pairs
183          (B) TYPE: nucleic acid
184          (C) STRANDEDNESS: single
185          (D) TOPOLOGY: linear
186      (ii) MOLECULE TYPE: DNA (genomic)
187      (vi) ORIGINAL SOURCE:
188          (A) ORGANISM: GPD2
189      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
190      GAATTCGAGC CTGAAGTGCT GATTACCTTC AGGTAGACTT CATCTTGACC CATCAACCCC      60
191      AGCGTCAATC CTGCAAATAC ACCACCCAGC AGCACTAGGA TGATAGAGAT AATATAGTAC      120
192      GTGGTAACGC TTGCCTCATC ACCTACGCTA TGGCCGGAAT CGGCAACATC CCTAGAATTG      180
193      AGTACGTGTG ATCCGGATAA CAACGGCAGT GAATATATCT TCGGTATCGT AAAGATGTGA      240
194      TATAAGATGA TGTATACCCA ATGAGGAGCG CCTGATCGTG ACCTAGACCT TAGTGGCAAA      300

```

RAW SEQUENCE LISTING

DATE: 01/09/2003

PATENT APPLICATION: US/09/308,207

TIME: 11:14:33

Input Set : N:\vernette\US09308207.raw

Output Set: N:\CRF4\01092003\I308207.raw

```

195 AACGACATAT CTATTATAGT GGGGAGAGTT TCGTGCAAAT AACAGACGCA GCAGCAAGTA 360
196 ACTGTGACGA TATCAACTCT TTTTTTATTA TGTAATAAGC AAACAAGCAC GAATGGGGAA 420
197 AGCCTATGTG CAATCACCAA GGTCGTCCCT TTTTCCCAT TTGCTAATTT AGAATTTAAA 480
198 GAAACCAAAA GAATGAAGAA AGAAAACAAA TACTAGCCCT AACCTGACT TCGTTTCTAT 540
199 GATAATACCC TGCTTTAATG AACGGTATGC CTTAGGGTAT ATCTCACTCT GTACGTTACA 600
200 AACTCCGGTT ATTTTATCGG AACATCCGAG CACCCGCGCC TTCTCAACC CAGGCACCGC 660
201 CCCAGGTAAC CGTGCGCGAT GAGCTAATCC TGAGCCATCA CCCACCCAC CCGTTGATGA 720
202 CAGCAATTCG GGAGGGCGAA AATAAACTG GAGCAAGGAA TTACCATCAC CGTCACCATC 780
203 ACCATCATAT CGCCTTAGCC TCTAGCCATA GCCATCATGC AAGCGTGAT CTCTAAGAT 840
204 TCAGTCATCA TCATTACCGA GTTTGTTTTT CTTCACATGA TGAAGAAGGT TTGAGTATGC 900
205 TCGAAACAAT AAGACGACGA TGGCTCTGCC ATTGGTTATA TTACGCTTTT GCGGCGAGGT 960
206 GCCGATGGGT TGCTGAGGGG AAGAGTGTTT AGCTTACGGA CCTATTGCCA TTGTTATTCC 1020
207 GATTAATCTA TTGTTACGCA GCTCTTCTCT ACCCTGTCAT TCTAGTATTT TTTTTTTTTT 1080
208 TTTTGGTTT TACTTTTTTT TCTCTTGCC TTTTCTTCTT GTTACTTTTT TTTACTTTTT 1140
209 TTTTCTTCC ACTAAGCTTT TTCTTGATT TATCCTTGGG TTCTTCTTTC TACTCCTTTA 1200
210 GATTTTTTTT TTATATATTA ATTTTAAAGT TTATGTATTT TGGTAGATTC AATTCTCTTT 1260
211 CCCTTTCCTT TTCCTTCGCT CCCCTTCCTT ATCAATGCTT GCTGTCAGAA GATTAACAAG 1320
212 ATACACATTC CTTAAGCGAA CGCATCCGGT GTTATATACT CGTCGTGCAT ATAAAATTTT 1380
213 GCCTTCAAGA TCTACTTTC TAAGAAGATC ATTATTACAA ACACAACTGC ACTCAAAGAT 1440
214 GACTGCTCAT ACTAATATCA AACAGCACAA AACTGTGCAT GAGGACCATC CTATCAGAAG 1500
215 ATCGGACTCT GCCGTGTCAA TTGTACATTT GAAACGTGCG CCCTTCAAGG TTACAGTGAT 1560
216 TGGTCTGGT AACTGGGGGA CCACCATCGC CAAAGTCATT GCGGAAAACA CAGAATTGCA 1620
217 TTCCCATATC TTCGAGCCAG AGGTGAGAAT GTGGGTTTTT GATGAAAAGA TCGGCGACGA 1680
218 AAATCTGACG GATATCATAA ATACAAGACA CCAGAACGTT AAATATCTAC CCAATATTGA 1740
219 CCTGCCCCAT AATCTAGTGG CCGATCCTGA TCTTTTACAC TCCATCAAGG GTGCTGACAT 1800
220 CCTTGTTTTT AACATCCCTC ATCAATTTTT ACCAAACATA GTCAAACAAT TGCAAGGCCA 1860
221 CGTGGCCCTT CATGTAAGGG CCATCTCGTG TCTAAAAGGG TTCGAGTTGG GCTCCAAGGG 1920
222 TGTGCAATTG CTATCCTCCT ATGTTACTGA TGAGTTAGGA ATCCAATGTG GCGCACTATC 1980
223 TGGTGCAAAAC TTGGCACCGG AAGTGGCCAA GGAGCATTGG TCCGAAACCA CCGTGGCTTA 2040
224 CCAACTACCA AAGGATTATC AAGGTGATGG CAAGGATGTA GATCATAAGA TTTTGAAATT 2100
225 GCTGTTCCAC AGACCTTACT TCCACGTCAA TGTCATCGAT GATGTTGCTG GTATATCCAT 2160
226 TGCCGGTGCC TTGAAGAACG TCGTGGCACT TGCATGTGGT TTCGTAGAAG GTATGGGATG 2220
227 GGGTAACAAT GCCTCCGCAG CCATTCAAAG GCTGGGTTTA GGTGAAATTA TCAAGTTCGG 2280
228 TAGAATGTTT TTCCAGAAT CCAAAGTCGA GACCTACTAT CAAGAATCCG CTGGTGTTGC 2340
229 AGATCTGATC ACCACCTGCT CAGGCGGTAG AAACGTCAAG GTTGCCACAT ACATGGCCAA 2400
230 GACCGGTAAG TCAGCCTTGG AAGCAGAAAA GGAATTGCTT AACGGTCAAT CCGCCAAGG 2460
231 GATAATCACA TGCAGAGAAG TTCACGAGTG GCTACAAACA TGTGAGTTGA CCAAGAATT 2520
232 CCCAATTATT CGAGGCAGTC TACCAGATAG TCTACAACAA CGTCCGCATG GAAGACCTAC 2580
233 CGGAGATGAT TGAAGAGCTA GACATCGATG ACGAATAGAC ACTCTCCCCC CCCCTCCCCC 2640
234 TCTGATCTTT CCTGTTGCCT CTTTTTCCCC CAACCAATTT ATCATTATAC ACAAGTTCTA 2700
235 CAACTACTAC TAGTAACATT ACTACAGTTA TTATAATTTT CTATTCTCTT TTTCTTTAAG 2760
236 AATCTATCAT TAACGTTAAT TTCTATATAT ACATAACTAC CATTATACAC GCTATTATCG 2820
237 TTTACATATC ACATCACCGT TAATGAAAGA TACGACACCC TGTACACTAA CACAATTTAA 2880
238 TAATCGCCAT AACCTTTTCT GTTATCTATA GCCCTTAAAG CTGTTTCTTC GAGCTTTTCA 2940
239 CTGCAG

```

(2) INFORMATION FOR SEQ ID NO: 7:.

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3178 base pairs

(B) TYPE: nucleic acid

VERIFICATION SUMMARY

DATE: 01/09/2003

PATENT APPLICATION: US/09/308,207

TIME: 11:14:34

Input Set : N:\vernette\US09308207.raw

Output Set: N:\CRF4\01092003\I308207.raw

L:20 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:21 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:1773 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=59
L:1858 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=60
L:1885 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=61
L:1912 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=62
L:1937 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=63
L:1962 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=64
L:1993 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=65
L:2063 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=67
L:2181 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:68